Evolved Digital Circuits and Genome Complexity

Morten Hartmann, Per Kristian Lehre and Pauline Haddow

Department of Computer and Information Science
Norwegian University of Science and Technology
{mortehar,lehre,pauline}@idi.ntnu.no

The 2005 NASA/DoD Conference on Evolvable Hardware, June 29 - July 1, Washington DC, USA

Outline

- Introduction
  - Motivation
  - Artificial Development
  - Complexity
- Experiments
  - Evolution of Circuits
  - Encoding Circuits
  - Measuring Complexity
- Results
  - Measured Complexity
- Conclusions
  - Conclusions and future work

The 2005 NASA/DoD Conference on Evolvable Hardware, June 29 - July 1, Washington DC, USA

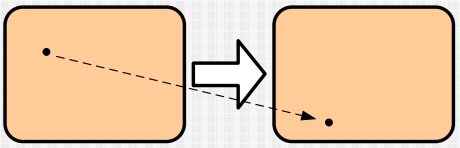
Motivation

- Scalability issues in EHW
  - Limited to relatively small systems
  - Digital circuits, typically 3-4 bit multipliers and adders
  - Some ~100+ gates
  - Each new element in the genotype adds a dimension to the search space
- Limits the size and complexity achievable today through evolutionary approaches
- Possible improvements
  - Hierarchical, Divide-and-conquer etc.
  - Tuning of representation and EA

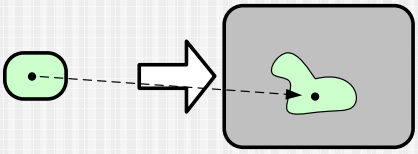
Artificial Development

- Using direct mapping, one need large genotypes to describe large phenotypes
- The human genome consists of about 3 billion base pairs (10^9)
- The human brain is believed to contain about 10^{14}-10^{15} synapses
- The process of biological development creates a complex creature from a (seemingly) less complex piece of information
- An indirect developmental mapping from genotype from phenotype
- This papers viewpoint with regards to scalability
  - Also self-organization, self-repair, gene regulatory networks etc.

Artificial Development



Direct mapping



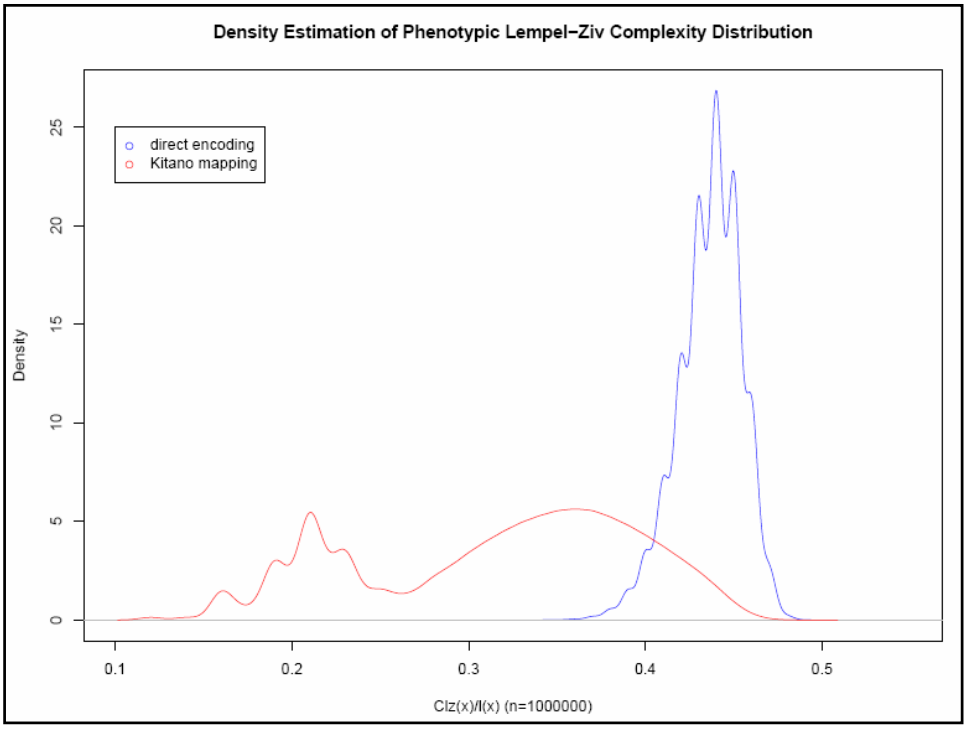
Indirect mapping

(deterministic)

Space of genotypes

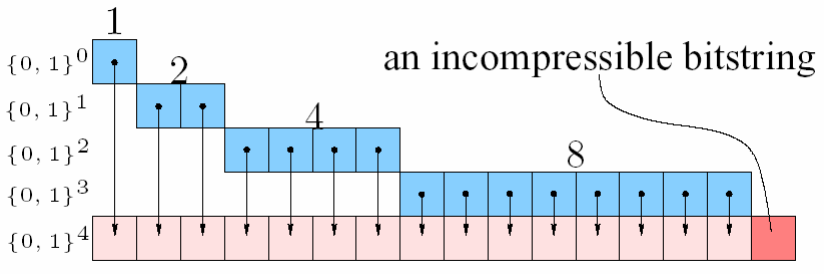
Complexity

- Want large and complex designs
  - Not just growth
- Measuring complexity (Lehre)
  - Algorithmic complexity (Kolmogorov)
  - Length of the shortest program that can output the string
- This complexity is incomputable
  - Approximation; Lempel-Ziv compression
  - Commonly used compression (zip)
  - LZ-complexity of a string is the length of its compressed version
- Measure the complexity of evolved circuits
  - Artificial development is a kind of decompression
  - Measuring complexity by compression indicates how small genotypes can be used to represent the relevant phenotypes
  - Are they suited for artificial development? (with regards to complexity)



Compressibility

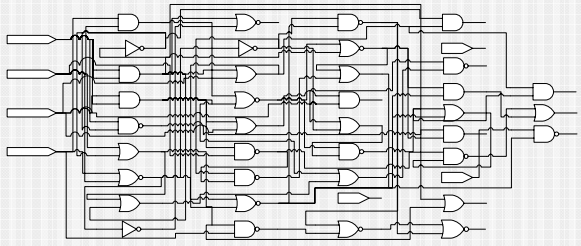
- Incompressibility lemma
  - Most binary strings are relatively incompressible



Outline

- Introduction
  - Motivation
  - Artificial Development
  - Complexity
- Experiments
  - Evolution of Circuits
  - Encoding Circuits
  - Measuring Complexity
- Results
  - Measured Complexity
- Conclusions
  - Conclusions and future work

Evolution of Circuits

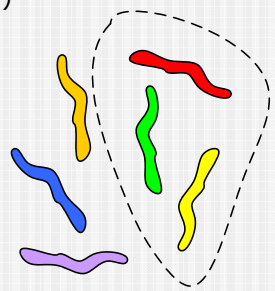


Genotype

Phenotype

Evolutionary Algorithm

- Cartesian Genetic Programming (J. Miller)
- Combinatorial circuits
  - 2-bit multiplier (200)
  - 2-bit adder (200)
- Evolutionary Algorithm
  - Tournament Selection (3/20)
  - Mutation rate 5%
  - Crossover rate 20%
- Genotype of 40 gates
  - NOR, NAND, OR, AND, NOT, VCC, GND
- Neutral elitism



Encoding Circuits

- Need a binary representation to measure LZ-complexity
- Encode circuits
  - Straight-forward
  - One-Hot
  - Unused-to-Zero

Encoding Circuits

	SYMBOLIC			STRAIGHT			UNUSED-TO-ZERO				ONE-HOT		
label	type	in A	in B	type	in A	in B	type	active	in A	in B	type	in A	in B
0													
1													
2													
3													
4	AND	3	1	011	11	011	011	y	11	01	0001000	11	011
5	NOT	4		010	100	000	000	n	000	000	0010000	100	00000
6	OR	0	2	101	000	010	101	y	000	010	0000010	000	010
7	NOR	2	1	110	010	001	110	y	010	001	0000001	010	001
38	VCC			001	000000	000000	001	y	000000	000000	0100000	000000	000000
39	GND			000	000000	000000	000	n	000000	000000	1000000	000000	000000
40	NAND	30	22	100	011110	010110	000	n	000000	000000	0000100	011110	010110
41	AND	15	27	011	001111	011011	011	y	001111	011011	0001000	001111	011011
42	OR	37	34	101	100101	100010	101	y	100101	100010	0000010	100101	100010
43	NAND	38	24	100	100110	011000	100	y	100110	011000	0000100	100110	011000

Measuring Complexity

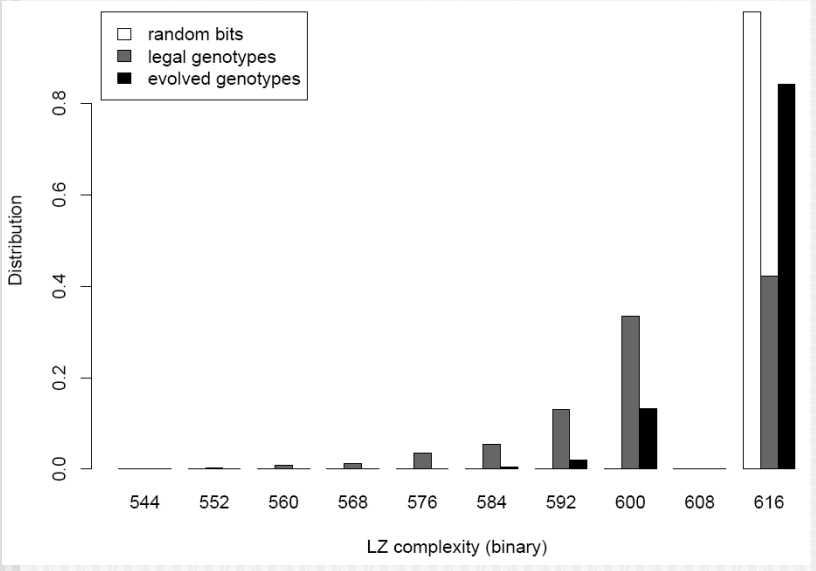
- LZ-complexity
  - Compress using **`zlib`**
  - Compresses to bytes
    -> 8-bit granularity
- Measure the complexity of 200 evolved multipliers and adders
- Compare complexity of circuits to:
  - Random binary strings of the same length (10k)
  - Randomly selected legal genotypes (10k)
    - Search space complexity

Outline

- Introduction
  - Motivation
  - Artificial Development
  - Complexity
- Experiments
  - Evolution of Circuits
  - Encoding Circuits
  - Measuring Complexity
- Results
  - Measured Complexity
- Conclusions
  - Conclusions and future work

Measured Complexity

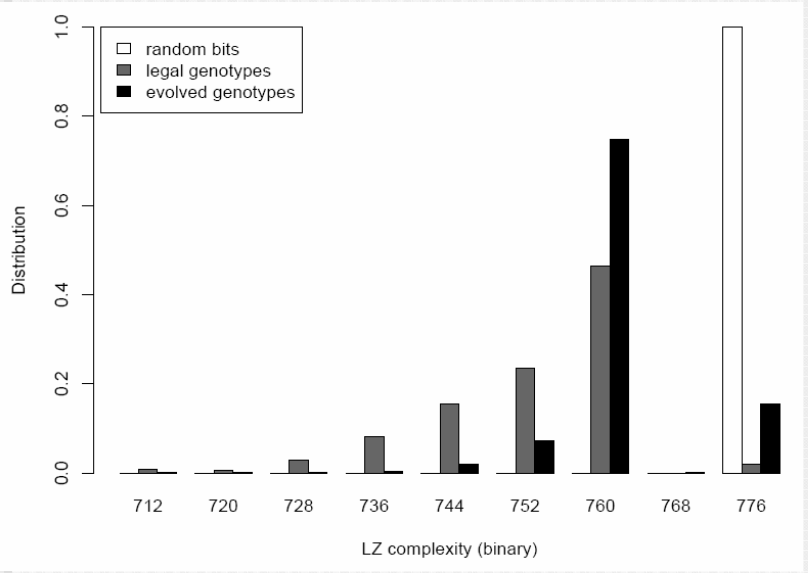
Straight forward encoding



The 2005 NASA/DoD Conference on Evolvable Hardware, June 29 - July 1, Washington DC, USA

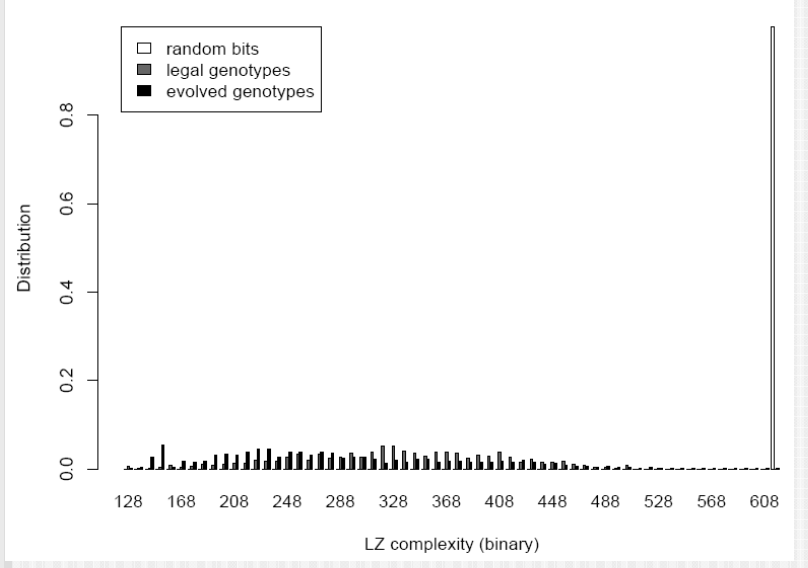
Measured Complexity

One-hot encoding

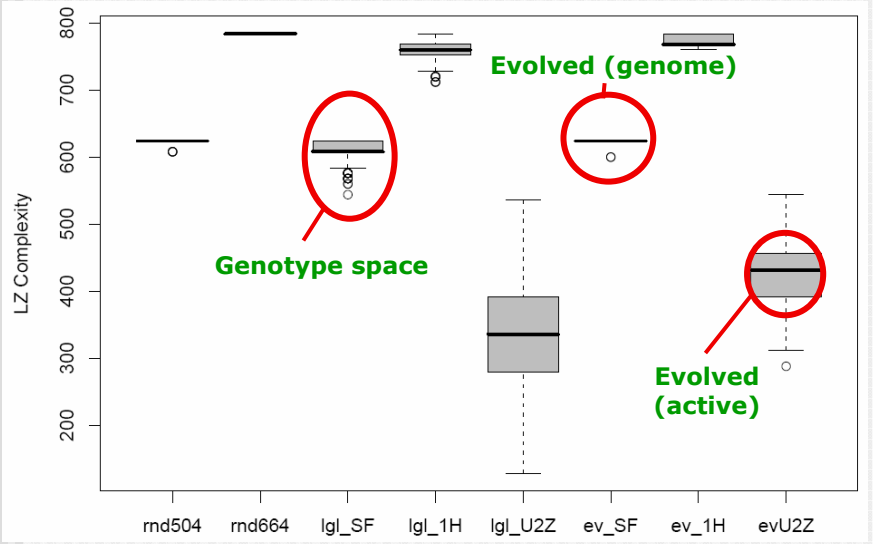


Measured Complexity

Unused-to-zero encoding



Measured Complexity



Outline

- Introduction
  - Motivation
  - Artificial Development
  - Complexity
- Experiments
  - Evolution of Circuits
  - Encoding Circuits
  - Measuring Complexity
- Results
  - Measured Complexity
- Conclusions
  - Conclusions and future work

Conclusions

- Implemented a measure for genotypic complexity of digital circuits
- Potential of reducing the search space by a huge factor… in theory
- LZ-decompression could act as an indirect mapping
- But…
  - Large genotypes are not necessarily bad
  - Representation and mapping features
  - Search space structure
  - Distance-correlation

Thank You !